JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1987, 20, 83–88 NUMBER 1 (SPRING 1987)

COMPUTER-ASSISTED OBSERVER TRAINING

ROGER F. BASS

UNIVERSITY OF WISCONSIN—MADISON

Computerized interactive videotapes were used to train college students to use a 10-second partial-interval observational recording system. Students viewed videotapes and scored response occurrences on a computer keyboard. Incorrect scoring resulted in immediate computerized feedback and rescoring.

Two independent variables, three versus seven target behaviors, and maintenance (thinning feedback) versus nonmaintenance (no feedback) training were compared. Average accuracy values across all target behaviors were consistently above 90%, and the lowest accuracy values obtained for most target responses typically met current standards for interobserver agreement. Data indicate the standard practice of occasional observer agreement checks may be inadequate.

DESCRIPTORS: observational technology, interactive video, computers

Direct observation techniques continue to be the most used data collection procedures cited in the *Journal of Applied Behavior Analysis* (Bass & Aserlind, 1984; Kelly, 1977) despite a decade of research that catalogs problems with training observers to use them (e.g., Farkas & Tharp, 1980; Mash & Makohonuik, 1975; Reid, 1970; Romanczyk, Kent, Diament, & O'Leary, 1973). This literature demonstrates the need for an observer training technology that consistently establishes and maintains accurate observational repertoires.

To address this need, microcomputer and video technologies were merged to create a computer-assisted observer training procedure that uses interactive video to train interval observation. Validation of this approach addressed these two questions: (a) Does computer-assisted observer training establish accurate observational repertoires, and (b) are these repertoires maintained?

METHOD

Observer Trainees

Twelve junior and senior level college students from the University of Wisconsin–Madison served as observer trainees. All were naive to the nature and purpose of interval observation techniques.

Apparatus

Hardware. The interactive video apparatus consisted of a 256K IBM-PC, a Tecmar VCR Controller Board, a Sony SL-2700 VCR, and a 13-inch color television. The VCR controller board encoded and read frame numbers on beta video and controlled VCR functions such as fast forward, stop, and reverse. (The Tecmar Corporation no longer supports the interface used in this research. A less expensive interface with even greater capabilities that works with home quality VCRs is available from Softward Assistance Corporation, 520 University Ave., Madison, Wisconsin 53703. Two alternative hardware systems designed for broadcast quality VCRs include (a) the BCD 450 VTR Control Interface manufactured by Video Associates Labs, 3933 Steck Ave., B-106, Austin, Texas 78759, and (b) TCR-3500 Time Code Reader plus TCG 3200 Timecode Generator, both built by FOR-A Company, LTD, 49 Lexington Street, West Newton, Massachusetts 02165.)

Videotapes. Target behaviors were depicted in 6-min videotapes of five special education students and their teacher during an industrial arts class. Rates of target behaviors varied from nearly 0% to 100% on each tape.

Videotapes were scored with a program that allowed repeated viewing of each interval. Later,

The author thanks Dale Liebert for his technical assistance.

Reprints may be obtained from Roger F. Bass, Department of Counseling and Special Education, University of Akron, Akron, Ohio 44325.

the computer assessed observational accuracy by comparing a trainer's and trainee's scorings.

The term *accuracy* is distinguished from *agreement* by denoting that observational records are compared to a standard endowed with more validity. For example, in this study, interval data collected in real time were compared to interval data collected by reviewing a permanent product record (i.e., a videotape). Agreement values compare observational records obtained under the same conditions.

Target Behaviors

Student target behaviors included *handling materials,* that is, obtaining, moving, or cleaning work-related materials; *student talk,* that is, any student verbal behavior including singing; and *on-task,* that is, any student behavior consistent with the completion of a workshop task.

Teacher target behaviors included both *verbal and nonverbal praise*; *giving instructions,* for example, modeling or discussing correct behavior, asking questions, and giving commands; *stopping a behavior,* that is, any verbal or nonverbal attempt to stop the occurrence of some ongoing response; and *touching hair,* that is, contact between the teacher's hand and hair.

Procedure

Observer training was carried out via several computer programs. The program *Target Behavior Recall* trained observers to identify written examples of targeted events. Scoring 90% correct on a target behavior quiz advanced the observer to the next training step. These sessions terminated with the observer receiving a $2.00 voucher that could be redeemed at the end of the study.

A program called *Establishing an Observational Repertoire* trained observers to identify videotaped examples of targeted events. Instructions covered how to score responses and the correction procedure. This procedure included (a) indicating that an error had occurred, (b) rescoring the interval, and (c) presenting the next interval if rescoring was correct, or otherwise naming the missed target responses and then reviewing the interval before proceeding.

Following each session for establishing an observational repertoire, the experimenter and observer calculated observational accuracy using this occurrence/nonoccurrence agreement formula: agreements on occurrences and nonoccurrences divided by agreements on occurrences and nonoccurrences plus disagreements on occurrences and nonoccurrences.

Training to establish an observational repertoire consisted of three steps with increasingly greater complexity values (i.e., the product of the number of target behaviors times the number of individuals observed). Complexity values for three target behaviors were 4, 5, and 6 whereas complexity values for seven target behaviors were 5, 10, and 16. Observers progressed one step when 90% accuracy was obtained across all target responses or when accuracy stabilized. Step 3 terminated when accuracy stabilized or was at least 90% across all target behaviors for 3 consecutive days. Observers then began either a maintenance or nonmaintenance condition.

Maintenance training involved three modifications of the previous phase. First, end-of-session feedback regarding observational accuracy was terminated. Second, the percentage of intervals checked for accuracy was thinned from 50% to 20% to 5.55%. Checked intervals were randomly selected by the computer. The criterion for changing feedback levels was stability or 90% accuracy across all target behaviors. Finally, a monetary consequence followed the accurate scoring of checked intervals. This consequence equaled $5.00 divided by the number of checked intervals and was added to the observers' usual $2.00 compensation. Therefore the highest amount that could be earned was constant at $7.00 per session. The maintenance phase ended when data stabilized or 90% accuracy was achieved across all target behaviors for 5 consecutive days.

The nonmaintenance condition involved terminating all feedback. Observers received an additional $5.00 per session and were told that the

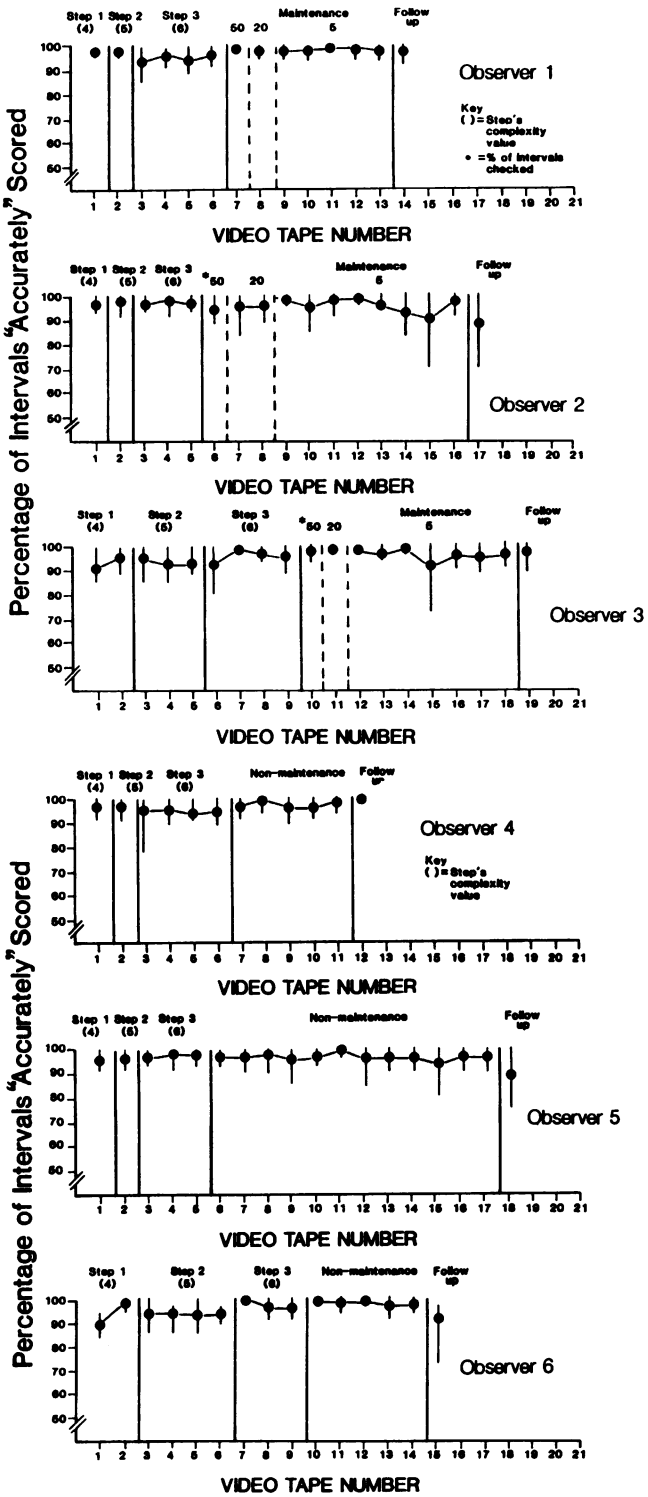

Figure 1. Range and average accuracy of observers scoring three target behaviors.

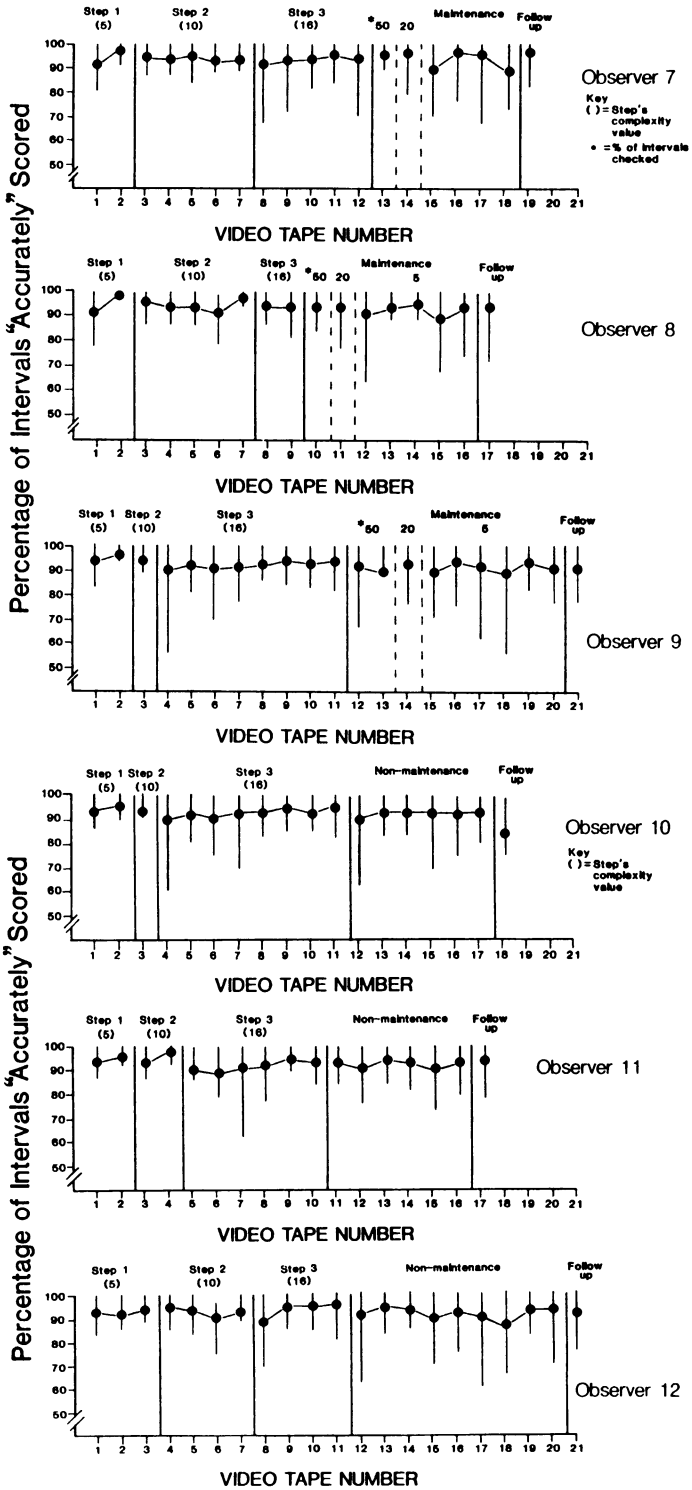

Figure 2. Range and average accuracy of observers scoring seven target behaviors.

experimenter would occasionally spot check their observational records. The criterion for the termination of this condition was stability or at least 90% accuracy across target behaviors for five consecutive sessions.

Observers completed a maximum of three sessions per day. The minimum intersession interval was the duration of the last session. The maximum interval was 72 hours.

RESULTS

All observers passed the target behavior recall test by the third session.

Establishing an Observational Repertoire

Three-target-behavior observers completed steps one and two in five to nine videotapes. Criteria for step 3 were met within four tapes. All inter- and intraobserver average accuracy fell within 11 percentage points across all tapes and none dropped below 89%. The ranges of observational accuracy fell below 80% only once. The high average accuracy indicates that the variability resulted from one or a few outliers.

Seven-target-behavior observers required between 9 and 12 videotapes to complete steps 1–3. Observational variability typically increased during the first few tapes of step 3. Again, consistently high average accuracy indicates that variability was due to outliers.

Maintenance and Nonmaintenance Conditions

Maintenance and nonmaintenance average accuracy for the seven-target-behavior observers approximated values obtained during the training condition. Variability was slightly greater during maintenance and nonmaintenance, but when it did increase, as on tapes 12, 15, 17, and 18, it did so across more than one observer and returned to previously obtained accuracy ranges. This indicates that the tape, and not a deteriorating observational repertoire, caused the lowered accuracy values. Analysis of the seven-target-behavior observers' data indicates that between 50% and 89% of their errors involved misscoring the target behavior *student talk*. In all cases, average accuracy was consistently above 90% in all conditions.

Maintenance and nonmaintenance data both across and within the three-target-behavior observers reflect little variability in either range or average accuracy. The greatest variability was specific to tape 15.

Follow-up Data

Follow-up data consistently fell within previously obtained ranges of average accuracy and variability. When this was not the case (e.g., observers 2, 5, and 6), the extreme variability was consistent with other observers' performance on that tape.

DISCUSSION

Computer-assisted observer training produced observational repertoires that accurately reflected a wide range of targeted events. Maintenance and nonmaintenance training generated equivalent degrees of observational accuracy. Inaccurate observations were concentrated on particular tapes and the target behavior *student talk*. Because observers made the same kinds of errors, their agreement remained high even though their accuracy fell. This illustrates the need for checks on observer accuracy, not agreement.

In summary, computer-assisted observer training can establish accurate observational repertoires, be modified to fit trainees' needs, provide exacting control for researching observational repertoires, and simplify data collection from videotapes.

Future research should assess the relative efficacy and cost effectiveness of computer-assisted observer training and traditional observer training procedures.

REFERENCES

Bass, R. F., & Aserlind, L. (1984). Interval and time-sample data collection procedures: Methodological issues. In K. Gadow (Ed.), *Advances in learning and behavioral disabilities: Vol. 3* (pp. 1–39). Greenwich, CT: JAI Press.

Farkas, G. M., & Tharp, R. G. (1980). Observation pro-

cedure, observer gender, and behavior valence as determinants of sampling error in a behavior assessment analogue. *Journal of Applied Behavior Analysis,* **13,** 529–536.

Kelly, M. B. (1977). A review of observational data-collection and reliability procedures reported in the *Journal of Applied Behavior Analysis. Journal of Applied Behavior Analysis,* **10,** 97–101.

Mash, E. J., & Makohonuik, G. (1975). The effects of prior information and behavioral predictability on observer accuracy. *Child Development,* **46,** 513–519.

Reid, J. (1970). Reliability assessment of observation data: A possible methodological problem. *Child Development,* **41,** 1143–1150.

Romanczyk, R. G., Kent, R. N., Diament, C., & O'Leary, K. D. (1973). Measuring the reliability of observational data: A reactive process. *Journal of Applied Behavior Analysis,* **6,** 175–184.

Received January 2, 1986
Initial editorial decision April 16, 1986
Revisions received July 19, 1986; August 19, 1986; November 17, 1986
Final acceptance November 19, 1986
Action Editor, Ron Van Houten